

California State University, San Bernardino

CSUSB ScholarWorks

Theses Digitization Project

John M. Pfau Library

1996

Do levels of self-monitoring, self-efficacy, and anxiety moderate the relationship between perception of dimensions and performance in assessment centers?

Jane Kathryn Paget

Follow this and additional works at: <https://scholarworks.lib.csusb.edu/etd-project>



Part of the [Psychology Commons](#)

Recommended Citation

Paget, Jane Kathryn, "Do levels of self-monitoring, self-efficacy, and anxiety moderate the relationship between perception of dimensions and performance in assessment centers?" (1996). *Theses Digitization Project*. 1157.

<https://scholarworks.lib.csusb.edu/etd-project/1157>

This Thesis is brought to you for free and open access by the John M. Pfau Library at CSUSB ScholarWorks. It has been accepted for inclusion in Theses Digitization Project by an authorized administrator of CSUSB ScholarWorks. For more information, please contact scholarworks@csusb.edu.

DO LEVELS OF SELF-MONITORING, SELF-EFFICACY, AND ANXIETY
MODERATE THE RELATIONSHIP BETWEEN PERCEPTION OF DIMENSIONS
AND PERFORMANCE IN ASSESSMENT CENTERS?

A Thesis
Presented to the
Faculty of
California State University,
San Bernardino

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
in
Psychology: Industrial/Organizational

by
Kathryn Jane Paget

June 1996

DO LEVELS OF SELF-MONITORING, SELF-EFFICACY, AND ANXIETY
MODERATE THE RELATIONSHIP BETWEEN PERCEPTION OF DIMENSIONS
AND PERFORMANCE IN ASSESSMENT CENTERS?

A Thesis
Presented to the
Faculty of
California State University,
San Bernardino

by

Kathryn Jane Paget

June 1996

Approved by:


Janet Kottke, Chair, Psychology


Janelle Gilbert


Diane Pfahler

6/6/96
Date

ABSTRACT

Many organizations use a selection technique known as an Assessment Center to identify potential employees. This technique commonly involves the use of four or five exercises typically an in-basket, role play, oral presentation, and a leaderless group discussion to assess a potential employee's ability to perform a given job. The popularity of assessment centers has risen from their predictive validity and utility. The lack of construct validity in assessment centers, however, has always been problematic. This study investigated the possibility that ability to perform well in an assessment center is influenced by ability to correctly perceive the dimensions assessed in a particular exercise. In addition, it was hypothesized that the relationship between dimensions identified and increased performance would be moderated by self-monitoring, self-efficacy, and anxiety. Although the results did not support the initial hypotheses, they do suggest new avenues of research to explore regarding variance in performance in assessment centers. For example, the purpose of a particular assessment center combined with the value of the outcome to the individual may help explain the lack of construct validity in assessment centers.

ACKNOWLEDGMENTS

The completion of this project has only been possible through the help and assistance of a great many people.

In particular, I would like to thank my thesis advisor, Dr. Janet Kottke not only for her guidance, assistance, and support throughout the thesis process, but also for her continuous support of my academic and professional growth over the last two years.

I would also like to especially thank Dr. Diane Pfahler, for her invaluable assistance in the collection of data for this project and her unwavering moral support, encouragement, and feedback.

Many thanks also go to Dr. Janelle Gilbert for her support not only during my thesis project but also in providing me with the opportunity to expand my skills into the realm of statistics teaching.

I am forever indebted to Bernadette Babasa, Maryann Christie, Cheryl Simmons, Leslie Haile, Jennifer Mersmen, Rudy Sanchez, Carrie Snow, and Ming Yu for taking time from their own theses to assist with my project. The completion of this thesis would have been impossible without their willing help and assistance.

Many thanks also go to Steve Waldman, Tom Richards, Carmen Carswell, and all the Media Services staff for their cheerfulness, moral support, and technical assistance.

I would also like to thank Nancy Pfefferle, Senior Personnel Analyst at County of San Bernardino, Human Resources for her flexibility in scheduling my work hours which allowed me to complete this project.

Finally, I would like to thank my husband Martin for his unwavering support, encouragement, and enthusiasm, not to mention hours of proof-reading, editing, and computer assistance. His unfailing wit, humor, and belief in my ability has been instrumental not only in the completion of this project but in my college success over the last five years.

TABLE OF CONTENTS

ABSTRACT.....	iii
ACKNOWLEDGMENTS.....	iv
LIST OF TABLES.....	vii
LITERATURE REVIEW.....	1
PURPOSE OF STUDY.....	16
HYPOTHESES.....	17
METHOD.....	18
Participants.....	18
Materials.....	19
Design and Procedure.....	24
RESULTS.....	25
Ancillary Results.....	26
DISCUSSION.....	28
Recommendations For Future Research.....	31
APPENDIX A: Demographic Information.....	35
APPENDIX B: Group Discussion Exercise.....	36
APPENDIX C: Self-Monitoring Questionnaire.....	38
APPENDIX D: Self-Efficacy Questionnaire.....	39
APPENDIX E: State-Anxiety Questionnaire.....	40
APPENDIX F: Individual Feedback Form.....	41
BIBLIOGRAPHY.....	42

LIST OF TABLES

Table 1. Factor Matrix: Self-Efficacy Scale.....	32
Table 2. Principal Components Analysis: Rating Dimension.....	33
Table 3. Correlation Matrix.....	34

LITERATURE REVIEW

Assessment centers are an increasingly popular method of identifying future managerial potential in organizations. More than 2000 organizations are currently using some type of assessment center program as a means of selection, placement, early identification of management potential, career management, or training (Gaugler, Rosenthal, Thornton, & Bentson 1987). Growth in the use of assessment centers as an aid in decision making has been seen in industrial, educational, military, government, and other organizational settings ("Guidelines and Ethical Considerations for Assessment Centers," 1989.)

The assessment center is typically organized around a set of dimensions based on aspects of behavior of successful managers. The dimensions assessed can include such qualities as leadership, planning, delegating, decisiveness, and communication skills. These skills are assessed by the use of a number of exercises which usually include an in-basket exercise, role play, oral presentation, and leaderless group discussion. Trained assessors rate each participant on his/her performance during the exercises, then confer to assign an overall rating for each participant on each of the various dimensions.

The premise of assessment centers is that they can improve selection of candidates for a particular job by the

use of exercises designed to simulate situations experienced in that job. The popularity of this selection technique has risen in part from its predictive validity. In one of the earliest studies using an assessment center, the Management Progress Study, conducted by the American Telephone and Telegraph Company, found that one of the most significant findings over the course of the 20 year study was that success as a manager was highly predictable. Success was defined as advancing further than the modal person in one's cohort (p. 183). Based on this definition, it was found that of the college graduates judged most promising by the assessment center, 43% reached a level of management higher than that of their modal peer, compared to only 20% of those judged less favorably at assessment (Bray, 1982). These results are particularly significant because the assessment center ratings were available only to the researchers in the study. As such, promotion decisions were made without any knowledge of the assessment center ratings. These results suggest that individual characteristics as measured by the assessment center were important determinants of managerial success.

The use of assessment centers as an aid in management selection was also demonstrated by Campbell and Bray (1993) in a follow-up to the original AT & T study. Their results found that 51% of those predicted to become middle managers did in fact make this level of management, compared to only

14% of those who were not predicted to make it into middle management but subsequently succeeded in doing so.

Similarly, a study by Moses and Boehm (1975), in a further follow-up to the original AT & T study which correlated current management level at the end of 1973 for 4846 women assessed between 1963 and 1971, suggested that assessment center performance was strongly related to subsequent promotions into management and advancement within management for women. For example, of the participants judged to be more than acceptable or acceptable in terms of advancement potential, 71% and 70% were subsequently promoted. In comparison, only 53% and 31% of those judged to be questionable and not acceptable, respectively, were subsequently promoted. This study, therefore, demonstrated that the assessment center predicted future performance for women as accurately as it did for men.

Hinrichs (1978), also investigated the predictive validity of an assessment center. Forty-seven individuals assessed in a management assessment center were followed up eight years later to determine the predictive validity of the assessment center. The results suggest that overall assessment center rating and a general management evaluation of individual potential were significantly related, with an adjusted multiple correlation of .58, to position level attained after eight years for the 30

individuals still with the company.

Perhaps even more persuasive evidence for the predictive validity of assessment centers is provided by a meta-analysis of 50 assessment center studies containing 107 validity coefficients by Gaugler et al (1987). These studies were drawn from published and unpublished reports, included experimental studies in which there was no feedback to participants, studies that compared the subsequent performance of assessed and non-assessed groups, correlational studies with feedback to assesses, and concurrent validation studies. After correcting for sampling error, restriction of range, and criterion unreliability, the mean validity coefficient was .37.

The improvement in predictive validity of assessment centers over traditional selection techniques has also been evaluated in terms of utility. As defined by Cascio and Silbey (1979), "Utility analysis, in brief, is the determination of expected institutional gain or loss anticipated to result from various courses of action or outcomes" (p. 108). In their Monte Carlo study they evaluated the utility of an assessment center by the use of the Brogden-Cronbach-Gleser continuous variable utility model. After defining their cost assumptions, they varied six parameters, such as the validity and cost of the assessment center, the selection ratio, and the standard deviation of the criterion. These hypothetical results

suggest that "even assessment centers with validities as low as .10 showed positive gains in utility over random selection."

In a comparative utility analysis, Cascio and Ramos (1986) indicated that even under the most conservative assumptions, the assessment center is cost-effective in comparison with an interview process. They estimated that the dollar gain in improved performance by using assessment centers instead of the interview process to select first level managers was \$2,676 per selectee per year in terms of improved performance.

In a similar study assessing the utility of assessment centers, Hogan and Zenke (1986) evaluated four selection procedures for hiring school principals. An interview, assessment center, selected assessment center exercises, and paper and pencil inventories were compared in terms of validity, processing costs, and dollar gain expected from the implementation of the procedure. The assessment center consisted of performance tests, interviews, and paper and pencil inventories. The selected assessment center exercises were a streamlined assessment center, the components of which were "based on conceptual contribution to job performance and empirical validity" (p 938). Using a pool of 115 individuals who applied for seven positions, the selected assessment center exercises and assessment center had the greatest utility with an expected dollar

value performance gain in excess of \$58,000 and \$43,000 respectively over the traditional interview alternative.

As the previous research demonstrates, there is a great deal of support for the predictive validity and utility of assessment centers. In contrast, the research explaining how assessment centers work is equivocal and problematic.

Assessment centers are purported to be successful in predicting future performance because they provide raters with an opportunity to infer personal qualities and traits that have been determined through job analysis to be relevant to success (Byham, 1980). However, the evidence for the construct validity of the dimensions used in assessment centers is discouraging. For example, a number of studies have revealed the lack of construct validity of assessment center ratings. In essence, ratings on the same dimension across different exercises do not correlate as highly as do ratings on different dimensions in a single exercise. For example, Sackett and Dreher (1982) examined the interrelationships among ratings between and within exercises in three assessment centers. In all three organizations the ratings within each exercise correlated more highly than across exercise ratings of the same dimension. The authors suggest that assessment ratings do not, in fact, measure the intended constructs. They suggest that possible explanations for the low across-

exercise correlations are 1) that behavior in an assessment center is situationally determined, therefore consistency across exercises should not be expected, and 2) that differences are due to low inter-rater reliability (p. 406).

Similar questions concerning what is being measured by assessment centers have also been raised by Harris, Becker, and Smith (1993). In an examination of two different assessment center scoring methods, 1) rating each dimension across all exercises before rating an overall exercise, and 2) rating each exercise independently, they found that dimension ratings correlated more highly with different dimensions in the same exercise than with ratings of the same dimension in other exercises. Ultimately, the switching to the nontraditional method of scoring management tasks, namely the within-dimension method as opposed to the within-exercise method, did not alleviate this problem.

Various explanations have been put forward as to why assessment centers appear to lack construct validity. Cohen and Sands (1978), for example, hypothesized that varying the order of presentation of exercises would differentially influence participant's performance during an assessment centers (p. 38). Five traditional assessment center exercises were used; in-basket, leadership problem, a problem solving exercise, and two leaderless group

discussions with a total of 67 government service managers who were randomly assigned to one of four order presentations. However, the results demonstrated that controlling the order of the exercises produced no significant differences between the four groups.

It has also been hypothesized that the number of assessment center dimensions is a determinant of assessor accuracy in rating performance on an assessment center (Gaugler & Thornton, 1989). They found that raters who rated a small number of dimensions classified behaviors more accurately and made more accurate ratings than did raters who rated a large number of dimensions. They suggest that the job of an assessor in an assessment center is exceedingly complex and the more demands that are made on assessors, the more difficult it is for an assessor to discriminate between the different dimensions that are being assessed.

Highhouse and Harris (1993) suggested that the different exercises within an assessment center constitute different situations, and that individual behavior varies from situation to situation. They therefore proposed that this variance of behavior across situations may explain individual differential performance on assessment center exercises. Using a Q-sort method, 6 experienced assessors were asked to classify 25 assessment center performance constructs, such as "generates enthusiasm" and "maintains

composure", according to how descriptive the items were of the ideal candidate in each assessment center exercise. The degree of exercise similarity was then compared to archival data of candidates' performance in the assessment center. The results suggested that there was some relationship between assessor perceptions of exercise similarity and the consistency of candidate performance for a group discussion and scheduling exercise. That is, these exercises were perceived as being dissimilar by the raters and candidate performance was also inconsistent across these exercises. However, the assessors also perceived the group discussion and a fact-finding exercise to be dissimilar in terms of performance demands, yet candidate performance in these exercises was relatively consistent (p. 150). In general, assessors perceived each exercise situation to have little resemblance to the other exercises. This dissimilarity across situations may explain why behavior is not consistent across situations. Mischel (1968) notes that "behavior depends on the stimulus situations and is specific to the situation....The more dissimilar the evoking situations, the less likely they are to produce similar or consistent responses from the same individual" (p. 177).

If the exercises in an assessment center constitute different situations and individual behavior varies across situations, then the apparent lack of construct validity in

assessment centers may be explained by the demand characteristics of each exercise and an individual's ability to react appropriately to the situational cues provided by that exercise. As such, there may be individual characteristics that either facilitate or interfere with performance in an assessment center.

This idea was supported by a longitudinal study of participants' characteristics as well as their reactions to an assessment center conducted by Fletcher (1991). He found that there were consistent differences between successful and unsuccessful candidates, both before and after the assessment center, that were not attributable to the impact of the assessment decision. Unsuccessful candidates had significantly lower scores than the successful candidates on work ethic, inquisitiveness, status, aspiration, mastery, and depressed mood. The authors proposed that these results suggest that the unsuccessful candidate can be visualized as someone who does not find reward in working hard, or mastering difficult problems and who is less positive in mood and therefore less likely to shine in an assessment center or similar working environment.

An extension of this concept was proposed by Kleinmann (1993), who suggested that one of the differences between successful and unsuccessful candidates in an assessment center is the ability to identify the dimensions being

assessed. In his study he assessed the relationship between transparency of rating dimensions and performance in an assessment center. The premise behind this was that participants who were better able to identify what was being assessed in an assessment center would perform better. Fifty-six students participated in an assessment center. Following the completion of each exercise, participants were asked about their perceptions of each exercise. The number of dimensions recognized was significantly correlated with the overall rating score, $r=.30$ ($p < .05$). The results suggest that people who can more accurately identify dimensions perform better than those who cannot, indicating that there may be more to success in an assessment center than an objective measure of task-specific abilities.

Given that performance in an assessment center may be influenced by the ability to perceive what is being assessed in an assessment center, a logical extension of this concept is to investigate individual characteristics that may facilitate or interfere with the ability to provide the expected behavior once the demand characteristics of the situation have been identified. For example, Snyder's theory of self-monitoring suggests that high self-monitors are able to modify their behavior according to role prescriptions and other situational demands regardless of whether or not the exhibited behavior

is congruent with inner feelings, attitudes, or emotions. In contrast, low self-monitors are relatively insensitive to situation cues and are somewhat ineffective actors of roles that do not have high congruence with internal states (Snyder, 1979). For example, in a study of a large insurance company, an association was found between self-monitoring and job level. Employees who were managers and supervisors were typically high self-monitors while technical, clerical, and support staff were found to be low self-monitors. This outcome suggests that high self-monitors are more able to adapt to the situation and act as directed by the situation which facilitates their roles as supervisors or managers. (Snyder, 1987).

Based on Snyder's theory, it is expected that high self-monitors will perform better in an assessment center as compared to low self-monitors provided that they have identified the abilities being assessed. Anderson and Thacker (1985) found some support for the idea that self-monitoring would be related to success in an assessment center used in the selection of sales staff. A significant correlation was found between self-monitoring and overall assessment rating for women but not men. The authors explain these results by the hypothesis that impression management skills (self-monitoring) would be more crucial for women than men in a traditionally male-oriented organization.

In addition to self-monitoring ability, it is also possible that levels of self-efficacy will influence individual performance in an assessment center. Bandura (1991) states that, "Among the mechanisms of personal agency, none is more central or pervasive than people's beliefs about their capabilities to exercise control over their own level of functioning and over events that affect their lives" (p 257). An individual's belief in his or her own efficacy may influence choices made, the amount of effort used in a given endeavor, the degree of perseverance, whether thought patterns are self-hindering or self-aiding, and the amount of stress experienced in coping with taxing environmental demands (p. 257).

Bouffard-Bouchard (1989) examined the effect that levels of self-efficacy would have on a cognitive performance task. Performance was assessed by using indicators such as persistence, (the number of problems that a student worked on until a correct response was found), success, (the number of correct responses), and the level of certainty of the correctness of responses. The results of the study suggest that students in a high-efficacy group completed a significantly greater number of problems than did the low efficacy group. In addition, perceived self-efficacy was related both to task persistence and to ability to evaluate the correctness of responses.

In a comparable study that aimed to assess the specific effect that self-efficacy would have on actual academic performance, Mone (1994), found similar results. Students were asked to indicate their level of confidence for attaining each of three grade categories on their next examination. Performance was measured as the actual grade for the examination. The results suggest that outcome self-efficacy, as this scenario was defined by Mone, significantly predicted performance.

The influence of self-efficacy on willingness to participate in a public performance situation would seem to be particularly relevant in assessing the impact of self-efficacy on assessment center performance. Arch (1992) investigated whether willingness to participate in a presentation to a large audience on a topic that they knew well would be influenced by task efficacy, cognitive control efficacy (ability to control negative thoughts and worries), and affective control efficacy (ability to handle nervousness). The resulting correlations between willingness to participate and these different aspects of efficacy are .61, .60 and .59 respectively. These results build on previous research that suggests that self-efficacy may influence task-performance, and also suggest that these results may generalize to willingness to perform in a public domain. Applying these results to performance in an assessment center, which can be considered a public

performance situation, would suggest that self-efficacy may be not only be related to willingness to perform in a public situation, but also to subsequent performance in that situation.

Levels of anxiety may also influence performance in an assessment center. For example, Glass, Arnkoff, Wood, Meyerhoff, Smith, Oleshansky, & Hedges 1995, studied the effect that anxiety would have on performance in a career-related oral examination. Participants completed the State questionnaire from the State-Trait Anxiety-Inventory a total of six times, seven and two days before, immediately before and immediately after, and two and seven days after the examination. The results suggest that "...a number of state anxiety questionnaires were related to performance, so that the more anxiety reported by participants (and the closer these ratings were to the day of the examination), the poorer their actual performance" (p. 50.). The correlations between the state-anxiety questionnaire immediately before and immediately after the exam, and exam board performance were $-.43$ and $-.42$ respectively. An interesting aspect of this study was that general trait measures of anxiety did not predict board scores, implying that the relationship between anxiety and performance was based on anxiety concerning the specific task and not due to general levels of anxiety. Although, Glass et al. (1995), did not address the issue of the inverted-U theory

of anxiety and performance, the fact that this relationship was not found in this study may be due to the possibility that anxiety was not manipulated and, therefore, restriction of range may explain the linear relationship of these results.

In summary, it has already been demonstrated that ability to perceive the dimensions being assessed in an assessment center is positively correlated with performance in an assessment center. What has not yet been investigated is the effect that self-efficacy, self-monitoring, and anxiety may have on either facilitating or interfering with an individual's ability to act on his or her perceptions.

PURPOSE OF STUDY

The purpose of this study was to investigate individual characteristics that may lead to differential performance across exercises in an assessment center. Specifically, this study aimed to measure the extent to which performance on an assessment center exercise, the leaderless group discussion, is moderated by identification of the dimensions being assessed and levels of self-monitoring, self-efficacy, and anxiety.

HYPOTHESES

Hypothesis 1.

As suggested by the results of Kleinmann's study (1993), the ability to identify the constructs being assessed in an assessment center exercise will be positively related to subsequent performance in that exercise.

Hypothesis 2.

The relationship between the ability to identify the constructs being assessed in an assessment center exercise and performance on that exercise, will be moderated by self-monitoring. It is not expected that gender differences will be found in this study as it is not expected that participation in a group discussion exercise will be seen as more typically male or female oriented. Therefore, the purpose of including this measure in this study is to investigate whether Anderson and Thacker's (1985) results are replicated, and if there is any empirical support for applying Snyder's self-monitoring theory as a moderator of performance in an assessment center.

Hypothesis 3.

The results from Hypothesis 1 will be moderated by self-efficacy. As supported by Bandura's theory and Arch's (1992) study investigating the willingness to participate in a public presentation, it is hypothesized that self-

efficacy for behavior in a group discussion will be positively correlated with ratings on the assessment center exercise.

Hypothesis 4.

The results from Hypothesis 1 will be moderated by levels of perceived anxiety. As suggested by the study conducted by Glass et al. (1995), performance will be negatively impacted by increased levels of state anxiety.

METHOD

Participants

Eighty-six California State University, San Bernardino students participated in the study. Based on Cohen (1992), this number was deemed sufficient to assess the possibility of a medium effect at the $p < .05$ level for a multiple regression procedure with 4 predictors. The sample consisted of 18 males and 68 females, whose ages ranged from 18 to 58 (mean = 26.12; sd = 8.18). The class level breakdown of the sample was as follows: Freshman, 10 (11.6%); Sophomores, 14 (16.1%); Juniors, 26 (29.9%); Seniors, 34 (39.1%); and Graduate Students, 2 (2.3%). Ethnicity was distributed as follows Asian, 12 (14%); African American, 11 (12.8%); Caucasian, 39 (45.3%); Hispanic, 18 (20.9%); and Other, 6 (7%). Of the eighty-six participants, nine had previously participated in an Assessment Center. Two of these previous experiences with an assessment center were for selection purposes, two for

developmental feedback, three were in connection with an experiment, and two participants failed to provide any additional information. Possible differences between those who had participated in an assessment center previously and those that had not were assessed by a t-test but yielded no significant differences on average rating score, dimensions identified, self-efficacy, self-monitoring, nor anxiety.

Participants were offered extra credit for participating in the study. All participants were treated in accordance with the "Ethical principles of psychologists and code of conduct" (American Psychological Association, 1992).

Materials

Demographic information:

Demographic information was collected by the means of a Participant Information Sheet (see APPENDIX A).

Assessment Center Exercise:

The Leaderless Group Discussion exercise as detailed in APPENDIX B was used as stimulus material for the participants. This exercise was developed by Dr. Janet Kottke as an integral part of the Practicum Class for the Master of Science in Psychology: Industrial/Organizational degree at California State University, San Bernardino. The leaderless group discussion exercise was chosen as the stimulus material as it was thought that this exercise would produce the greatest variability in behavior, as it

assesses a number of dimensions --- interpersonal, presentation, and communication skills.

Self-Monitoring:

Self-monitoring was assessed by the Lennox and Wolfe (1984) revised version of Snyder's 25-item Self-Monitoring Scale (see APPENDIX C). This scale is comprised of two subscales, ability to modify self-presentation and sensitivity to expressive behavior of others. The alpha coefficients, reported by Lennox and Wolfe for each of these subscales is .77 and .70, with the alpha coefficient for the entire scale being .75. The alpha coefficients for self-presentation, sensitivity to expressive behavior, and the entire scale for this sample were similar: .76, .72, and .77 respectively.

Self-Efficacy:

Self-efficacy was assessed by using a modified version of Riggs, Warka, Babasa, Betancourt, & Hooker (1994) Personal Efficacy Beliefs Scale, (see APPENDIX D). The alpha reliability of the original scale was reported as .86. Validity coefficients for the original scale were reported as varying from .22 (performance), .25 (organizational commitment), and .30 (job satisfaction). As the original scale, developed by Riggs et al. was adapted for the purpose of assessing self-efficacy in a group-discussion exercise, a separate reliability analysis and factor analysis was conducted for this scale. The

factor analysis yielded two factors identified by Eigenvalues that were greater than 1 (see Table 1). Only one question (# 3) loaded on Factor 2; this item was removed from the scale, and all resulting analyses calculated for the nine remaining items. The alpha for this revised scale was .91, and significant correlations were obtained with self-monitoring ($r = .365$, $p < .001$, $r^2 = .13.$), pre anxiety ($r = -.456$, $p < .001$, $r^2 = .21$), and post anxiety scores ($r = -.378$, $p < .001$, $r^2 = .14.$), suggesting that the revised scale demonstrated adequate convergent and divergent validity.

Specific research in the area of self-efficacy has typically included task-specific measures of this construct. For example, in a study assessing the effects of self-efficacy on training outcomes, Mathieu, Martineau & Tannenbaum (1993) asked students "...to rate the extent to which they believed that they could score at least each of eight bowling scores;...that corresponded to the grade levels that they could earn" (p. 134). In a similar study, Mathieu and Button (1992) assessed the effect of self-efficacy on personal goals and performance. Self-efficacy was measured by asking subjects to rate the extent to which they believed that they could score a particular score on a performance task. In general, when self-efficacy has been measured in connection with a specific task, it has been found that task-specific scales are better predictors of

performance on cognitive tasks than general measures of self-efficacy (Wang & Richarde, 1988).

Anxiety:

Anxiety was assessed by the State Questionnaire from the State-Trait Anxiety Inventory (see APPENDIX E). This inventory assess anxiety levels that are specific to the situation. The A-State has demonstrated high internal consistency with Cronbach's alpha coefficients ranging from .83 to .92. The alpha coefficients for this sample were .89 for the pre-measure and .90 for the post measure.

Perception of Constructs Being Measured:

Participants' perception of the constructs being assessed by the Leaderless Group Discussion was assessed by giving them a checklist of 10 possible constructs, including 5 bogus items and the 5 true items (see APPENDIX F). Participants were asked to identify the 5 dimensions that they thought were being assessed by the Leaderless Group Discussion. Each participant's degree of perceptiveness as to what was being measured by the assessment center was assessed by the number of dimensions that were correctly identified. The alpha for the reliability among dimensions correctly selected (the true items) by participants was -.39 and the alpha for the reliability among dimensions incorrectly selected (the bogus items) by participants was -.47.

Rating Score:

Eight graduate students were recruited to rate the participants performance in the leaderless group discussion exercise. They participated in a one-hour training session using behavioral examples from a mock videotape on how to rate participants, using a frame of reference technique. The mock videotape portrayed four actors participating in the group discussion exercise. They exhibited behaviors at the high and low end of each dimension scale. Immediately after the watching the videotape, the raters formed into rating pairs and rated one of the actors in the videotape.

After practicing reaching a consensus rating, the scores were compared to the researcher's who had already rated each of the actors on their performance.

This method of rater training is supported by research by Athey and McIntyre (1987), that indicates that frame of reference training improved retention of training, improved accuracy, and less halo over information only training or no training at all. Sulsky and Day (1992) also found that frame of reference training led to better rating accuracy and better classification accuracy.

Each participant was assessed by two raters who reported a consensus score for each participant on each of the four dimensions assessed. These dimensions were listening and comprehension, consulting skills, sensitivity to others, and presentation style. On each of these

dimensions the minimum and maximum possible scores were 1 and 5 respectively. (Each participant's rating score was calculated by averaging their score across the four rating dimensions).

A principal components analysis was performed on the average rating score to determine whether there was justification for using an average score as opposed to treating each dimension as a separate variable. This analysis, (see Table 2) yielded only one significant factor. In addition to the principal components analysis, the alpha reliability of the rating scores was .88.

Design and Procedure

On arrival, participants were told the nature of the project and signed an informed consent sheet indicating their willingness to participate in the study. Participants were then asked to complete the general demographic information sheet, self-efficacy and self-monitoring questionnaires and the A-State anxiety scale. Participants were then asked to participate in the Leaderless Group Discussion exercise in groups of three or four. Their discussion was videotaped and evaluated at a later date. Upon completion of the discussion exercise, the A-State anxiety scale was re-administered as well as the checklist for participants to indicate what they thought was being assessed by the assessment center. After completion of the questionnaires, participants were

debriefed as to the nature of the study, and requested not to reveal its purposes to other participants.

RESULTS

Prior to any analyses being calculated, the data were examined for normality and linearity. All data demonstrated adequate variability. In addition, while the data were cleaned a random sample of 20% of the data was checked for data entry errors, with 100% of the data concerning rating scores and dimensions identified being checked as these variables were considered critical to the analyses.

Hypothesis 1. the ability to identify the constructs being assessed in an assessment center exercise will be positively related to subsequent performance in that exercise, was not supported. The correlation between dimensions identified and average rating for each participant was not significant ($r = -.032$) at $p < .05$.

Hypothesis 2. There was no significant correlation between self-monitoring and average rating score ($r = .073$) at $p < .05$. A t-test assessing the effect of gender on self-monitoring yielded no significant differences between males and females at $p < .05$.

Hypothesis 3. The correlation between self-efficacy and average rating score ($r = -.016$) was not significant at $p < .05$.

Hypothesis 4. Before analyzing this hypothesis, the anxiety data were tested for possible curvilinearity. No such relationship was evident, therefore, a Pearson's r was calculated in order to assess the relationship of this variable to performance on the group discussion exercise. There was no significant correlation between the anxiety pre-score and average rating score ($r = -.098$, ns) at $p < .05$.

A correlation matrix displaying correlations between each of the above variables is displayed in Table 3.

Ancillary Results

The following analyses were run to fully explore the data.

Average Rating

Correlations between each rating dimension, (listening and comprehension, consulting skills, sensitivity to others, and presentation style), and number of dimensions identified were calculated to assess whether the score on any one dimension was related to ability to identify the dimensions being assessed. However, none of these correlations were significant at $p < .05$.

To assess the possibility that raters were rating differentially, and that any significant effects were being averaged out, correlations were also calculated between the average rating score by each pair of raters and each variable of interest. However, none of these correlations were significant.

To determine whether performance on a particular dimension was different based on whether a participant successfully identified that dimension, five t-tests were calculated (one for each dimension). However, none of these were significant.

Self-efficacy

A correlation was run between age and self-efficacy to determine whether age may be acting as a moderator variable, however, this correlation was not significant at $p < .05$.

Gender

A t-test was calculated to assess whether there were differences in rating score based on gender, however this was not significant at $p < .05$.

Anxiety

A dependent t-test performed to assess whether there were differences between pre and post levels of anxiety was significant ($t(85) = 7.58, p < .001$). The mean anxiety score prior to the group discussion exercise was 1.95, while the mean anxiety score after the completion of the group discussion exercise was 1.65. Significant correlations were obtained between pre-anxiety scores and self-monitoring ($r = -.365, p < .01, r^2 = .13$), and self-efficacy ($r = -.453, p < .01, r^2 = .21$). Significant correlations were also obtained between post-anxiety scores and self-monitoring ($r = -.387, p < .05, r^2 = .15$), and

self-efficacy ($r = -.378$, $p < .01$, $r^2 = .14$). In addition, there was a significant negative correlation between the post anxiety score and average rating score ($r = -.220$, $p < .05$, $r^2 = .05$).

DISCUSSION

Although the hypotheses proposed by this study were not supported, the results obtained may still provide useful information in understanding the lack of construct validity in assessment centers.

Various analyses were conducted to ascertain whether there were methodological flaws in the data collection and in the use of the various scales. Reliability analyses of all scales used suggest that the data collected demonstrated acceptable reliability. In addition, significant correlations between certain variables provide support for the construct validity of these variables. For example, the significant positive correlations between self-efficacy and self-monitoring support the validity of the self-efficacy scale. Similar conclusions can be drawn from the negative relationships between self-efficacy and self-monitoring and anxiety. The significant difference between pre and post anxiety scores suggests that the lack of a significant negative relationship between anxiety and performance on the group discussion exercise is not due to a lack of apprehension concerning the exercise. However, it is possible that the limited range of responses for the

dimensions to be identified may have influenced participants' responses on their identification of dimensions being assessed. A more varied range of responses to choose from may have more accurately distinguished between those who clearly identified what was being assessed and what was not.

The unidimensionality of the consensus ratings were also evaluated by running a factor analysis, reliability analysis, and by comparing rating scores broken down by both dimension and by rating pair. The consistency of information suggests that rating inaccuracy did not affect the outcome of the study.

Given that the scales and rating scores seem to be providing accurate information, an explanation of the results may be in the behavior of the participants themselves. Initially, it was considered that participants had not taken the study seriously as their only incentive was extra credit. For example, the researcher had noticed that some participants did not appear to take the instructions seriously and aimed to complete the discussion exercise as quickly as possible. It was considered that these participants who did not seriously discuss the issues in coming to a consensus may be having a negative impact on the overall results. Therefore, it was decided to view a sample of the videotapes and re-calculate the analyses using only the subjects who appeared to take the exercise

seriously. A review of forty-six randomly selected subjects led to the identification of thirty subjects who appeared to participate in the group discussion exercise at a serious level. However, when the analyses were re-run based on this group of individuals there was no significant change in the results.

Individual performance may still, however, explain the results. Kleinmann, 1993, demonstrated that participants who were better able to identify the rating dimensions on a given exercise received higher ratings on those same exercises. In his study the participants took part in the assessment center from a job-applicant-training framework. Participants for Kleinmann's study were recruited by a student organization that regularly organized student seminars, and students paid a fee to attend the assessment center. In this scenario, where participants paid to participate in the study it is more likely that there was a greater incentive for students to perform at an optimum level. In the present study where there were no negative consequences for failing to perform at an optimum level, it is possible that participants did not perform in the same way that they would have done when participating in either a selection or developmental assessment center. As such, the results obtained in the present study suggest the level of performance that is observed in a group discussion exercise when there is no external incentive to perform.

The basis of the present study focused on participant's ability to identify the dimensions assessed by an assessment center exercise and then to measure their ability to match their behavior to what they perceived as the most desirable behavior. However, in order to expend this effort it seems logical that there would need to be some incentive for participants to behave in this way. Typical assessment centers used for either selection or developmental purposes would seem to provide more of this incentive.

Recommendations For Future Research

Based on these results it would seem appropriate to replicate the present study in both a developmental and selection setting. Such a study would help clarify the relationship between the purpose of the assessment center and subsequent performance. In such a replication it would also seem worthwhile to investigate the importance that participants place on the outcome of the assessment center as a variable of interest. It may be fallacious to assume participation in an assessment center for either selection or developmental purposes guarantees the participant's valence in the outcome. An examination of the situational variables and individual differences across a variety of situations may then help explain the lack of construct validity in assessment centers.

Table 1

Factor Matrix: Self-Efficacy Scale

SE Item	Factor 1	Factor 2
<hr/>		
SE01	<u>.81076</u>	-.30057
SE02	<u>.63933</u>	-.52077
SE03	.37870	<u>.72817</u>
SE04	<u>.82369</u>	.05161
SE05	<u>.82517</u>	-.14079
SE06	<u>.78527</u>	.09794
SE07	<u>.75874</u>	.20016
SE08	<u>.74600</u>	.27190
SE09	<u>.76800</u>	-.27595
SE10	<u>.74918</u>	.20669

Table 2

Principal Components Analysis: Rating Dimension

Initial Statistics:

Variable	Communality	Factor	Eigenvalue	Pct of Var	Cum Pct
RAT1	1.00000	* 1	2.99381	74.8	74.8
RAT2	1.00000	* 2	.51996	13.0	87.8
RAT3	1.00000	* 3	.26391	6.6	94.4
RAT4	1.00000	* 4	.22233	5.6	100.0

PC extracted 1 factor.

Factor Matrix:

	Factor 1
RAT1	.90364
RAT2	.90477
RAT3	.82462
RAT4	.82380

Table 3

Correlation Matrix**Correlations**

		average rating	mean self-efficacy score	mean pre-anxiety score	post anxiety average	no of dimensions identified	mean self-monitoring score
Pearson Correlation	average rating	1.000	-.016	-.098	-.220*	-.032	.073
	mean self-efficacy score	-.016	1.000	-.456**	-.378**	-.028	.365**
	mean pre-anxiety score	-.098	-.456**	1.000	.688**	.024	-.365**
	post anxiety average	-.220*	-.378**	.688**	1.000	-.063	-.387**
	no of dimensions identified	-.032	-.028	.024	-.063	1.000	-.143
	mean self-monitoring score	.073	.365**	-.365**	-.387**	-.143	1.000
Sig. (1-tailed)	average rating	.	.442	.184	.020	.385	.252
	mean self-efficacy score	.442	.	.000	.000	.397	.000
	mean pre-anxiety score	.184	.000	.	.000	.415	.000
	post anxiety average	.020	.000	.000	.	.281	.000
	no of dimensions identified	.385	.397	.415	.281	.	.095
	mean self-monitoring score	.252	.000	.000	.000	.095	.
N	average rating	87	86	86	87	87	86
	mean self-efficacy score	86	86	86	86	86	86
	mean pre-anxiety score	86	86	86	86	86	86
	post anxiety average	87	86	86	87	87	86
	no of dimensions identified	87	86	86	87	87	86
	mean self-monitoring score	86	86	86	86	86	86

*. Correlation is significant at the 0.05 level (1-tailed).

**. Correlation is significant at the 0.01 level (1-tailed).

APPENDIX A: Demographic Information

Participant Number: _____

Gender:

Male _____ Female _____

Age: _____

School Status:

Freshman _____ Sophomore _____ Junior _____ Senior _____

Grad Student _____

Ethnic Background:

Asian _____

African American _____

Caucasian _____

Hispanic _____

Native American _____

Other _____

Have you ever participated in an Assessment Center?

Yes _____ No _____

IF YES

How long ago did you participate in the assessment center? _____

What was the purpose of the assessment center (e.g. for selection or developmental purposes)

APPENDIX B: Group Discussion Exercise

Instructions

For this exercise, each of you will take on the role of a City of Bloomington Council Member. You will have 30 minutes to review the materials, discuss them, and come to a consensus as a group.

Problem Description

On March 5, the City Council of Bloomington received notification that Stanley and Sophie Kuchinski had willed their property at 125 Ridge Road to the city. The letter stated that the Kuchinskis had attached the following stipulations:

1. The Council must accept the donation within 3 months or forego any claim to it;
2. The Council must also decide on its use by this date;
3. If the Council chooses to lease, sell, or donate the property, it may do so to either a nonprofit or profit organization as long as the use "contributes to the quality of life of the community."

It is now June 3, and the Council members are meeting to make their decision. Prior to the meeting, they solicited requests and suggestions for use of the property, a brick structure located on approximately 1 acre of prime land.

The following "bids" were received:

1. The Friendship House, a United Way agency that runs programs for minority and disadvantaged youth, has requested that the city arrange a lease/purchase agreement. Bloomington is primarily a blue collar town with a growing black and Hispanic population. There is clearly a need for programs for disadvantaged young people, but neighbors in the area have vocally come out against the Friendship House proposal, crying, "We don't want 'them' over here."
2. Saint Stanislaus Church, whose property borders the Kuchinskis' on the east, has offered to buy the property at fair market value. The Church would tear down the building to create additional parking facilities to accommodate their growing crowds on Bingo nights. Many senior citizens support this use, since Bingo is one of the few recreational outlets for them, and "safe" off-street parking is at a premium.

3. A local builder has offered to buy the property and develop it into a moderately priced retirement condominium building. This would require special building permits, but would help the tax base and also be aesthetically pleasing.
4. A local women's group has proposed a three-year lease to set up a Women's Center. It would provide workshops, birth control and abortion counseling (and possibly a clinic), and also serve as a refuge for battered women. According to statistics compiled by the police, wife beating has increased drastically, probably as a result of downsizing which has caused layoffs in this working class community. The Rector of Saint Stanislaus' Church is strongly opposed to letting the women's group lease the property.
5. One of the major oil companies has submitted a bid several times fair market value for the property if the Council will grant a zoning change to allow a gas station. These funds could be used to buy sorely needed playground equipment but the gas stations would be an eyesore on Ridge Road.
6. John Lateck has offered to buy the property for \$300,000 and convert it to a "private club". He has assured the Council that it would not be an "ordinary" bar. Rumors are that John is a homosexual. Parents of children at Saint Stanislaus School have besieged the Council with letters smearing John and alleging that the property would turn into a gay bar.
7. Nafco, a statewide drug addiction service, has asked to lease the property to set up a drug rehabilitation center with residential facilities.

Which of these uses, if any, would you champion to the group?

The group MAY NOT vote to come to a conclusion; you must all agree to a single decision.

APPENDIX C: Self-Monitoring Questionnaire

Participant # _____

		CERTAINLY ALWAYS FALSE	GENERALLY FALSE	SOMEWHAT FALSE, BUT WITH EXCEPTION	SOMEWHAT TRUE, BUT WITH EXCEPTION	GENERALLY TRUE	CERTAINLY ALWAYS TRUE
1.	In social situations, I have the ability to alter my behavior if I feel that something else is called for.	0	1	2	3	4	5
2.	I have the ability to control the way I come across to people, depending on the impression I wish to give them.	0	1	2	3	4	5
3.	When I feel that the image I am portraying isn't working, I can readily change it to something that does.	0	1	2	3	4	5
4.	I have trouble changing my behavior to suit different people and different situations.	0	1	2	3	4	5
5.	I have found that I can adjust my behavior to meet the requirements of any situation I find myself in.	0	1	2	3	4	5
6.	Even when it might be to my advantage, I have difficulty putting up a good front.	0	1	2	3	4	5
7.	Once I know what the situation calls for, it's easy for me to regulate my actions accordingly	0	1	2	3	4	5
8.	I am often able to read people's true emotions correctly through their eyes.	0	1	2	3	4	5
9.	In conversations, I am sensitive to even the slightest change in facial expression of the person I'm conversing with.	0	1	2	3	4	5
10.	My powers of intuition are quite good when it comes to understanding others' emotions and motives.	0	1	2	3	4	5
11.	I can usually tell when others consider a joke to be in bad taste, even though they may laugh convincingly.	0	1	2	3	4	5
12.	I can usually tell when I've said something inappropriate by reading it in the listener's eyes.	0	1	2	3	4	5
13.	If someone is lying to me, I usually know it at once from that person's manner of expression.	0	1	2	3	4	5

APPENDIX D: Self-Efficacy Questionnaire

Participant # _____

Please think about your ability to do the tasks required by a group discussion. When answering the following questions, answer in reference to your personal experiences with group discussions.

	strongly disagree	disagree	disagree somewhat	agree somewhat	agree	strongly agree
1. I have confidence in my ability to express myself well in a group discussion.	1	2	3	4	5	6
2. I am an expert when participating in a group discussion.	1	2	3	4	5	6
3. When my performance is poor in a group discussion it is due to my lack of ability.	1	2	3	4	5	6
4. I doubt my ability to perform well in a group discussion.	1	2	3	4	5	6
5. I have all the skills needed to perform well in a group discussion.	1	2	3	4	5	6
6. Most people perform better in a group discussion than I do.	1	2	3	4	5	6
7. There are some aspects of participating in a group discussion that I cannot do well	1	2	3	4	5	6
8. The likelihood that I will do well in future group discussions is limited because of my lack of skills.	1	2	3	4	5	6
9. I am proud of my skills and abilities in a group discussion.	1	2	3	4	5	6
10. I feel nervous when others look at me while participating in group discussions.	1	2	3	4	5	6

APPENDIX E: State-Anxiety Questionnaire

Participant # _____

Directions: A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate number to the right of the statement to indicate how you feel *right now*, that is at *this moment*. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best.

	Not At All	Some What	Moderately So	Very Much So
1. I feel calm	1	2	3	4
2. I feel secure	1	2	3	4
3. I am tense	1	2	3	4
4. I am regretful	1	2	3	4
5. I feel at ease	1	2	3	4
6. I feel upset	1	2	3	4
7. I am presently worrying over possible misfortunes	1	2	3	4
8. I feel rested	1	2	3	4
9. I feel anxious	1	2	3	4
10. I feel comfortable	1	2	3	4
11. I feel self-confident	1	2	3	4
12. I feel nervous	1	2	3	4
13. I am jittery	1	2	3	4
14. I feel "high strung"	1	2	3	4
15. I am relaxed	1	2	3	4
16. I feel content	1	2	3	4
17. I am worried	1	2	3	4
18. I feel over-excited and rattled	1	2	3	4
19. I feel joyful	1	2	3	4
20. I feel pleasant	1	2	3	4

APPENDIX F: Individual Feedback Form

Participant # _____

Of the following ten group discussion skills, please check the **FIVE** that you believe were being rated in this exercise.

1. Ability to lead the group _____
2. Ability to explain issues to others. _____
3. Listening and Comprehension Skills _____
4. Ability to address important issues _____
5. Sensitivity to others needs _____
6. Leadership style _____
7. Cooperativeness and team-building _____
8. Directing the group process _____
9. Ability to persuade others of your point of view _____
10. Presentation style _____

BIBLIOGRAPHY

- American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. American Psychologist, 47, 1597-1611.
- Anderson, L.R. & Thacker, J. (1985). Self-monitoring and sex as related to assessment center ratings and job performance. Basic and Applied Psychology, 6 (4), 345-361.
- Athey, T.R. & McIntyre, R.M. (1987). Effect of rater training on rater accuracy: Levels-of-processing theory and social facilitation theory perspectives. Journal of Applied Psychology, 72 (4), 567-572.
- Bandura, A. (1991). Social theory of self-regulation. Organizational Behavior and Human Decision Processes, 50, 248-287.
- Bouffard-Bouchard, T. (1989). Influence of self-efficacy on performance in a cognitive task. Journal of Social Psychology, 130 (3), 353-363.
- Bray, D.W. (1982). The assessment center and the study of lives. American Psychologist, 37 (2), 180-189.
- Byham, W.C. (1982). Applying a systems approach to personnel activities. Training and Development Journal, Feb 1982, 86-90.
- Campbell, R.J. & Bray, D.W. (1993). Use of an assessment center as an aid in management selection. Personnel Psychology, 46, 691-699.
- Cohen, S.L. & Sands, L. (1978). The effects of order of exercise presentation on assessment center performance: One standardization concern. Personnel Psychology, 31, 35-45.
- Cascio, W.F. & Ramos, R.A. (1986). Development and application of a new method for assessing job performance in behavioral/economic terms. Journal of Applied Psychology, 71 (1), 20-28.
- Cascio, W.F. & Silbey, V. (1979). Utility of the assessment center as a selection device. Journal of Applied Psychology, 64 (2), 107-118.
- Cohen, J. (1992). A power primer. Psychological Bulletin, 112 (1), 155-159.
- Fletcher, C. (1991). Candidates' reactions to assessment center and their outcomes: A longitudinal study. Journal of Occupational Psychology, 64, 117-127.
- Gaugler, B.B., Rosenthal, D.B., Thornton III, G.C., & Bentson, C. (1987). Meta-analysis of assessment center validity. Journal of Applied Psychology Monograph, 72 (3), 496-511.

- Gaugler, B.B. & Thornton III, G.C. (1989). Number of assessment center dimensions as a determinant of assessor accuracy. Journal of Applied Psychology, 74 (4), 611-618.
- Glass, C.R., Arnkoff, D.B., Wood, H., Meyerhoff, J.L., Smith, H. R., Oleshansky, M.A., & Hedges, S.M. (1995). Cognition, anxiety, and performance on a career-related oral examination. Journal of Counseling Psychology, 42, 47-54.
- Guidelines and Ethical Considerations for Assessment Center Operations. (1989). Public Personnel Management, 18 (4), 457-470.
- Harris, M.M., Becker, A.S., & Smith, D.E. (1993). Does the assessment center scoring method affect the cross-situational consistency of ratings? Journal of Applied Psychology, 78 (4), 675-678.
- Highhouse, S. & Harris, M.M. (1993). The measurement of assessment center situations: Bem's template matching technique for examining exercise similarity. Journal of Applied Social Psychology, 23 (2), 140-155.
- Hinrichs, J.R. (1978). An eight year follow-up of a management assessment center. Journal of Applied Psychology, 63 (5), 596-601.
- Hogan, J. & Zenke, L.L. (1986). Dollar-value utility of alternative procedures for selecting school-principals. Educational and Psychological Measurement, 46, 935-945.
- Kleinmann, M. (1993). Are rating dimensions in assessment centers transparent for participants? Consequences for criterion and construct validity. Journal of Applied Psychology, 78 (6), 988-993.
- Mathieu, J. E. & Button, S. B. (1992). An examination of the relative impact of normative information and self-efficacy on personal goals and performance over time. Journal of Applied Social Psychology, 22, 1758-1775.
- Mathieu, J.E., Martineau, J. W., & Tannenbaum, S.I. (1993). Individual and situational influences on the development of self-efficacy: Implications for training effectiveness. Personnel Psychology, 46, 124-147.
- Mischel, W. (1968). Personality and assessment. New York: Wiley.
- Mone, M. A. (1994). Comparative validity of two measures of self-efficacy predicting academic goals and performance. Educational and Psychological Measurement, 54 (2), 516-529.

- Moses, J.L. & Boehm, V.R. (1975). Relationship of assessment-center performance to management progress of women. Journal of Applied Psychology, 60 (4), 527-529.
- Riggs, M.L., Warka, J., Babasa, B., Betancourt, R., & Hooker, S. (1994). Development and validation of self-efficacy and outcome expectancy scales for job-related applications. Educational and Psychological Measurement, 54, (3), 793-802.
- Russell, C.J. (1987). Person characteristics versus role congruency explanations for assessment center ratings. Academy of Management Journal, 30 (4), 817-826.
- Sackett, P.R. & Dreher, G.F. (1982). Constructs and assessment center dimensions: Some troubling empirical findings. Journal of Applied Psychology, 67 (4), 401-410.
- Snyder, M. (1979). Self-monitoring processes. In L. Berkowitz (ed.) Advances in experimental social psychology (Vol 12, pp. 86-128.) New York: Academic Press.
- Snyder, M. (1987). Public appearances-private realities: The psychology of self-monitoring (pp. 88-90). New York: W.H. Freeman and Company.
- Sulsky, L.M. and Day, D.V. (1992). Frame-of-reference training and cognitive categorization: An empirical investigation of rater memory issues. Journal of Applied Psychology, 77 (4), 501-510.
- Wang, A. Y., & Richarde, R.S. (1988). Global versus task specific measures of self-efficacy. The Psychological Record, 38, 533-541.